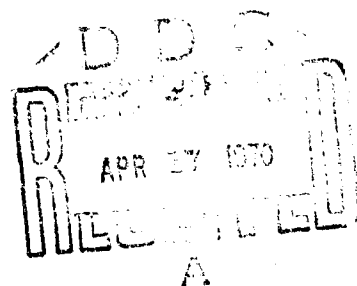# BRL

MEMORANDUM REPORT NO. 2025

A COMPUTER BASED SEARCHABLE FILE OF JOURNAL REFERENCES IN THE FIELD OF PARTIAL DIFFERENTIAL EQUATIONS

by

George C. Francis

February 1970

U.S. ARMY ABERDEEN RESEARCH AND DEVELOPMENT CENTER
BALLISTIC RESEARCH LABORATORIES
ABERDEEN PROVING GROUND, MARYLAND

B A L L I S T I C   R E S E A R C H   L A B O R A T O R I E S

MEMORANDUM REPORT NO. 2025

FEBRUARY 1970

A COMPUTER BASED SEARCHABLE FILE OF JOURNAL REFERENCES IN
THE FIELD OF PARTIAL DIFFERENTIAL EQUATIONS

George C. Francis

Applied Mathematics Division

A B E R D E E N   P R O V I N G   G R O U N D   M A R Y L A N D

BALLISTIC RESEARCH LABORATORIES

MEMORANDUM REPORT NO. 2025

GCFrancis/bj
Aberdeen Proving Ground, Md.
February 1970

A COMPUTER BASED SEARCHABLE FILE OF JOURNAL REFERENCES IN
THE FIELD OF PARTIAL DIFFERENTIAL EQUATIONS

ABSTRACT

The BRLESC computer has been programmed to prepare a searchable
and updatable magnetic tape file of journal references in the field of
partial differential equations. Multiple simultaneous queries
involving Boolean logic are permitted and have been used in automatic
indexing of the file prior to publication of a first version. Search
strategies are partially self-organizing. Examples of program use are
included in the report.

PRECEDING PAGE BLANK

TABLE OF CONTENTS

INTRODUCTION

A file of more than 4200 references to journal articles and some books in the field of partial differential equations, especially in numerical methods of solving them, has been accumulated. The BRLESC computer has been programmed to establish this file on magnetic tape, to update the file in various ways, to search the file for items of interest to a researcher, to edit the file for publication, and to index the file prior to publication.

This report describes some of the main features of the computer program and its use. A first edition of the file containing almost 3000 entries has been published using these techniques. [BRL Memorandum Report 1991 by J. H. Giese].

BACKGROUND

A subject of obvious importance to the Applied Mathematics Division as well as other laboratories at BRL is the solution of partial differential equations. Ready access to a central file of journal references in this field should prove of value to a researcher.

A mere tabulation of these references, sequenced by author, say, would be useful, but a facility to search for the occurrence of sets of terms of particular interest to a user would make the file more attractive to potential users. To this end it was decided to design and implement on the BRLESC computer a method of establishing, updating, searching, and indexing this file. Prepublication editing techniques were also to be considered. Magnetic tape was selected as the bulk storage medium for the file, as magnetic disks are not currently available on BRLESC.

A computer program for BRLESC 1 was designed by the author, written, tested and put to use. The original set of references, consisting of a few hundred entries, was punched on cards with the author(s) on the first

PRECEDING PAGE BLANK

7

card, marked uniquely to indicate "start of record." The "Initiate" mode of the program edited the records somewhat, assigned a sequential record number to each, and prepared the initial magnetic tape file.

Subsequent uses of the program in its "Update" mode have extended the file, deleted or replaced some records, modified certain records, renumbered the records, and thus prepared updated versions of the file on several occasions. The "Search" mode has been used to select references of interest to several users within BRL. Options provide for printing the entire reference or merely the record number, the latter being more useful for indexing purposes.

Indexing of a fixed "first volume" of the file (List A) was accomplished by a systematic computer search for some 150 sets of near-synonyms, usually fragments of words, in several natural languages. Some references were found to meet the criteria stated but not be pertinent, and these "false drops" were removed using a different option of the indexing program.

As an aid to publication an edited tape of List A in page format complete with page numbers and "page ejection" signals was prepared using another option of the BRLESC computer program. Similar paging was done for the index tape. The published result is now available as BRL Memorandum Report 1991.

THE PDE REFERENCES TAPE FILE

The current file of references to journal articles on numerical methods of solving partial differential equations (List A) is maintained on one partial reel of magnetic tape at BRL. There is a descriptive introduction at the beginning of the tape in order that a meaningful self-contained listing can be obtained by merely listing the tape. This introduction is followed by the main file entries (almost 3000 references), which in turn are followed by a sentinel or end-of-file entry.

The format of each of these parts influences the search and update programs; so an explanation of format is in order.

The introduction consists at present of a single line title, followed by two blank lines, a paragraph, six blank lines, a paragraph, two blank lines, a paragraph, and finally six more blank lines. All introduction lines are up to 80 characters in length. The usual computer input routines at BRL ignore single blank lines and thus drop them in copying (as in updating List A to make a new version on a new tape reel). Two consecutive blank lines do cause a special counter to be set in such a way that this condition can be tested and appropriate action taken to duplicate the two blank lines in the copied version. Hence, in the updating mode the introduction is read one line at a time, the counter tested, the line written on a new tape in the usual case and two blank lines written in the special case. This continues until the introduction is fully copied (recognized when the first author line, specially marked, is noted). In the search mode no duplicating of the introduction is necessary; so each line is read, tested for being an author line, and discarded until the first author line is found.

Reference entries are of two types, those with explicitly stated authors and those whose authors are omitted because identical to the authors of the preceding reference. The first line of each type can be recognized, however, and indeed must be if references are to be distinguished. Author lines have a special non-printing symbol as the first character of the line, and the surname of the author starts at the second character of that line. At present that special character is the equivalent of digits 5 and 8 double-punched in a single card column. This combination does not print on the BRL version of the IBM 1403 printer (i.e., prints a single blank character), although on a different BRL printer (Anelex) it appears as a colon. This character is recognizable by the computer, however.

9

The "second" line of each reference (or the first line if the author is omitted) has a reference sequence number in the first 7 character field with the title starting as character 8. This and any remaining lines of the reference entry are of length up to 87 characters each. The first seven characters of such remaining lines are blank. Thus the "second" line can be distinguished, as it must be if there is no author line.

Finally, the end of the file must be recognized, and for this purpose a sentinel, i.e. a dummy final record, is used. In particular, the sentinel used is a line with five commas as characters 1-5, suitable since no other line of the file has that form. (Many other sentinels could have been used instead.)

For completeness, typical entries are described below. This description is not essential to use of the file. See Figure 1.

:ALBASINY, E.L. / DAY, W.A.

    19. THE NUMERICAL SOLUTION OF LINEAR ONE-DIMENSIONAL
        PARABOLIC PROBLEMS BY THE METHOD OF MOMENTS.
        J. INST. MATH. AND ITS APPL. 4(1968), 140-162.
        REF. Z 1968, NO. 12B778.

:ALBRECHT, J.

    20. TAYLOR-ENTWICKLUNGEN UND FINITE AUSDRUECKE FUER
        (DELTA) U UND (DELTA 2) U.
        Z. ANGEW. MATH. MECH. 33(1953), 41-48. MR 14, P. 907.

    21. ZUM DIFFERENZVERFAHREN BEI PARABOLISCHEN
        DIFFERENTIALGLEICHUNGEN.
        Z. ANGEW. MATH. MECH. 37(1957), 202-212. MR 19, P. 462.
        REF. Z (1958), NO. 735.

:ALBRECHT, R. / URICH, W.

    22. A DIFFERENCE METHOD FOR THE APPROXIMATE SOLUTION OF
        THE INITIAL VALUE PROBLEM FOR SYSTEMS OF QUASILINEAR
        PARTIAL DIFFERENTIAL EQUATIONS OF FIRST ORDER.
        DEUX CONGR. ASSOC. FRANCAISE CALCUL ET TRAITEMENT
        INFORMATION (PARIS, 1961), 105-106. GAUTHIER-VILLARS,
        PARIS (1962). MR 31, NO. 2834.

Figure 1. Sample File Records

An author's name is given in the sequence: surname, comma, blank, initial, period, initial, period usually (certain oriental names are spelled out in full and considered surnames.) If there are several authors, all are listed on one line separated by the three characters blank, slash, blank. In general all entries for the same author (or set of coauthors) appear sequentially with the author stated explicitly only for the first (earliest) entry. Entries are alphabetical by surname of first author.

Included in a typical entry are title, bibliographic data on the journal in which the item appeared (volume, page, date, etc.), and usually one or more review journal entries. Usually the title is in its original language, but in some cases an English translation is given instead.

SEARCHING THE FILE

Since very little indexing (categorizing) of the entries in the file has been done, the usual search made is for the occurrence of any of a string of words, phrases, or fragments in any entry. A listing of the reference sequence numbers of the "winning" entries is the standard output. Certain other options can be used if desired.

Specifications for a search (or an update or other operation) are made using punched cards in a fixed format. (See Figure 2. for an example.) At least two cards are needed for a single search. The first has SEARCH starting at column 1 (and often nothing more). The second starts a list of equivalent search terms with SEARCH at column 1, TERMS at column 11, and a word or phrase or fragment at column 21. Such a "search term" may be up to 40 characters long at present (through column 60). If no additional terms are of interest, columns 79 and 80 are left blank. If equivalent terms are to follow, then OR is punched at columns 79-80 and additional terms are given at column 21 of the next card, continuing until no OR is required. Thus the last term of the set has columns 79-80 blank. Columns 1-20 are ignored after the first card of a set.

```
Col        Col        Col                                              Col
1          11         21                                               79

SEARCH
SEARCH     TERMS      THREEbDIM                                        OR
                      TRIDIM                                           OR
                      THREEbSPAC
```

Figure 2. Single Search Specification

The computer program then reads from the tape file one full record (bypassing the introduction), scans it for the occurrence anywhere, even in the middle of a word, of the first search term of the set. If a match is obtained, a "win" is noted and the reference sequence number is saved for later printing (or under certain options a suitable print-out is made immediately). If no match is obtained, the same record is searched for the next search term of the set, continuing until a match is obtained or no further terms are available.

After completing the processing of that record, the program reads the next record, processes it similarly, and continues until the end-of-tape sentinel is noted. At this point the tape file is rewound and any delayed printing accomplished. Next the program reads another set of search specifications, if any, and repeats the entire procedure.

Searching the entire file for one set of search terms at a time is time consuming; so where several independent sets of search terms are available, an option permits them to be processed together and then printed separately during the tape rewind. The number of such sets is entered at column 21 of the first specification card (the one with SEARCH at column 1 and blanks at column 11). Each set of terms has columns 79-80 blank in the last card and has SEARCH at column 1 and TERMS at column 11 of the first card of the set. (See Figure 3. for an example.). At present up to 30 sets of terms can be handled in one tape pass.

```
Col          Col          Col                                                    Col
1            11           21                                                     79

SEARCH                    3
SEARCH       TERMS        CONVERGE                                               OR
                          DIVERGE
SEARCH       TERMS        BOUNDARY                                               OR
                          GRENZSCH                                               OR
                          COUCHE
SEARCH       TERMS        CHARA                                                  OR
                          CARACT
```

Figure 3. Multiple Search Specifications

As a convenience to the user and to save computer time both blanks and hyphens are ignored during searches. Thus EIGENVALUE, EIGEN-VALUE, and EIGEN VALUE all match. Of course, EIGENWERT does not match any of these unless listed as an equivalent term. A search on the fragment EIGEN would find any of these, but would also match EIGENVECTOR, which might not be wanted. Some "false matches" are to be expected, however; so the output should be carefully checked for pertinence.

To date the current List A has been searched for about 150 sets of equivalent terms, and a partial index of List A is the result. Many references were found not to be pertinent and have been removed, using the computer itself to update the index tape. (See Figure 4.)

Recently the search capability has been extended to allow some use of the Boolean "and". Searches of the form T_1 and T_2 and ... and T_n are allowed, where the T_i may be individual terms or may be sets of equivalent terms connected by "ors" as in $(A_1$ or $A_2)$ and $(B_1$ or B_2 or $B_3)$. See Figure 5. for an example. Very limited use of this capability has been made to date.

No more general Boolean logic is available yet; but several independent searches can be made, and the resulting lists can mannually be compared, combined, and so forth. Addition of other logic will be considered in the future.

ENERGY METHOD

  1018 1379 1412 2123

ERROR BOUNDS

  144   187   215   312   392   405   406   413   827
  846 1191 1332 1400 1457 1493 1704 2074 2396
  2427 2431 2469

EULERIAN FORM

  842

FLUID DYNAMICS (SEE ALSO AERO-, GAS, AND HYDRO-DYNAMICS)

  167   998   999 1115 1202 1273 1444 2031 2513

Figure 4. Sample Index Entries

Col 1	Col 11	Col 21		Col 78
SEARCH				
SEARCH	TERMS	NUMER		bOR
		APPROX		AND
		DIFF		bOR
		DERIV		

Figure 5. Use of Boolean Logic

      Search terms and their associated Boolean connectives are kept in arrays, as are convenient links which permit bypassing terms which prove unnecessary. Suppose there are two requests in the form: t_1 and (t_2 or t_3) and (t_4 or t_5) and then u_1 and (u_2 or u_3 or u_4). A given record is first searched for t_1. If t_1 is not found, there is no need to search further; and the program directs control to the next request and a search for u_1. However, if t_1 is found, t_2 is sought next. Failure to find t_2 directs search to t_3, but success with t_2 allows bypassing t_3 and continuing with t_4. Failure with t_3 directs control to the next request and a search for u_1, whereas success with t_3 permits search for t_4. Failure with t_4 directs search to t_5, but success with t_4 permits bypassing t_5. This indicates success with the first request.

14
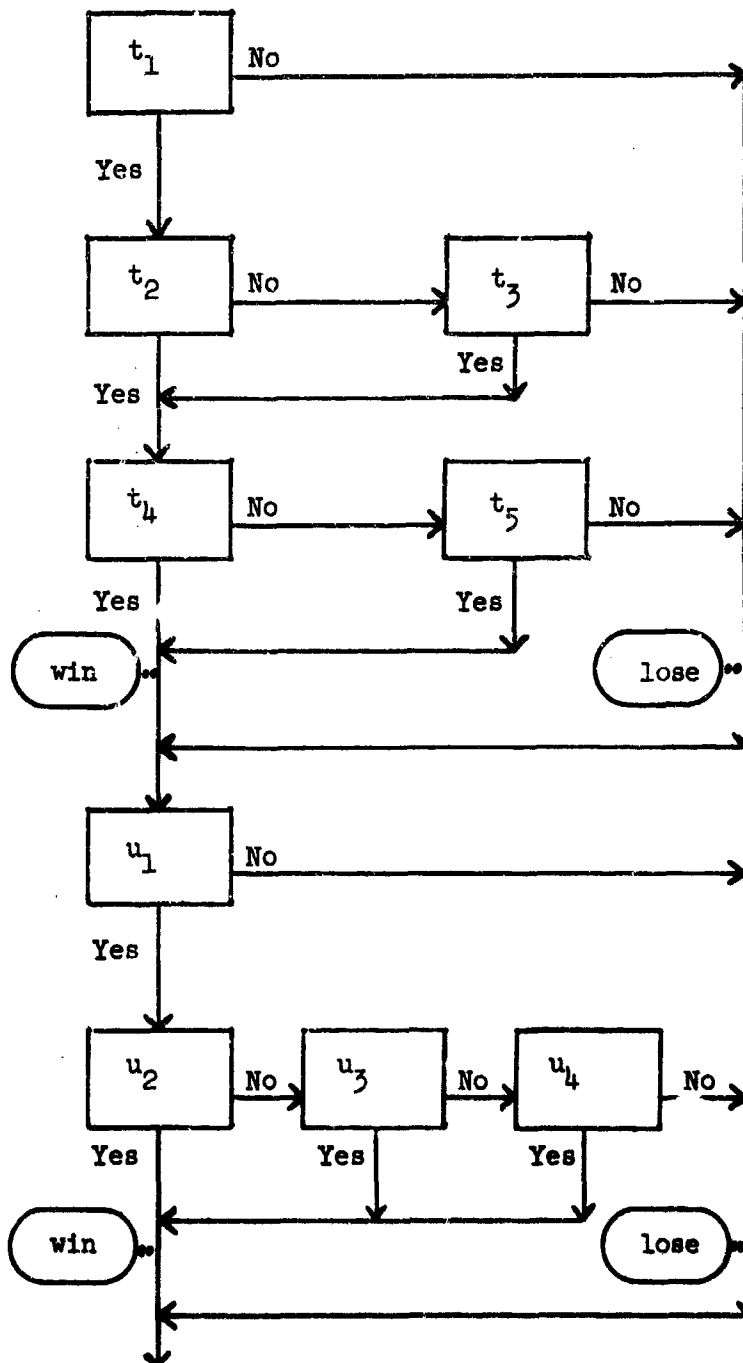
Request 1

Request 2

Request 3, if any

Figure 6. Self-Organizing Search Strategy

15

Then the second request is considered. Figure 6. indicates the above logic schematically. Each search is organized at the time the request is read from cards and used repeatedly.

Care by the user in stating the query can allow faster searches in some cases. For instance, u_1 and (u_2 or u_3 or u_4) is generally faster than (u_2 or u_3 or u_4) and u_1. In the first case failure with the first search (u_1) permits bypassing the rest, but in the second case failure with the first search (u_2) still requires further search for u_3 and at least one other term.

UPDATING THE FILE

Various options exist for changing the file (List A). Entire records can be inserted, incorrect or improperly sequenced records can be deleted in full, changes to part of a record can be made, and lines can be appended at the end of a record. In all of these, use is made of the sequence number of the record in question, or a near neighbor, as given in the latest version of the file. The mode in all these cases is UPDATE (as contrasted with SEARCH, etc). The specific types of updating are INSERTbAFTER, INSERTbBEFORE, DELETE, APPEND, and REPLACE.

To insert a new record (author, title, etc.) after a given record, say after number 1013 (and hence before 1014) a specification card with INSERTbAFTER in columns 1-12, 1013 at column 21, and blanks elsewhere is put ahead of the cards representing the record. Several records can be inserted after the same record, but each must have its own specification. The resulting sequence will be in the same order as the insertions.

To insert a new record before a specified record, the procedure is the same except that INSERTbBEFORE is punched in columns 1-13. Multiple records can be inserted before the same record. The resulting sequence will be in the same order as the insertions.

16

To delete a record, say 1025, punch DELETE in columns 1-6, 1025 at column 21, and blanks elsewhere. No record is required in this case, and the single card is sufficient.

To append something to the end of a specified record (such as an additional citation), punch APPEND in columns 1-6, the record numbers at column 21, and blanks elsewhere. Then starting at column 1 of a new card punch the exact information to be appended. There is an 80 column maximum to the amount appended at present, but several APPEND's can apply in succession to the same record. If the added information will fit into available space on the previously last line of the record, it is put there. Otherwise a new line is started.

REPLACE has several forms. For complete replacement of a given record by a new record, punch REPLACE at columns 1-7, the record number at column 21, and blanks elsewhere. Follow this with a complete record in the usual form (author, title, etc.) To replace just an author, say because of a misspelling, punch REPLACE at column 1, AUTHOR at column 11, and the record number at column 21; and on the next card starting at column 2 punch the correct version of the author line. (Column 1 of this card will be replaced by a 5-8 "punch".) (The other lines of the tape record will not be changed by this replacement.) If several successive records have the same (incorrect) author, only the first need be replaced. To replace one line other than the author line of a record, punch REPLACE at column 1, LINE at column 11, the record number at column 21, and the line number (1,2,3,...) at column 31. Line number 1 refers to the first (non-blank) line following the author line, 2 the next line, etc. Follow this card with a card giving the correct form of that line (neglecting the 7 column "indentation" or "record number" entry. A maximum of 80 characters is allowed. This causes a single line to be deleted and the substitute line inserted in its place.

Because a human error could cause the wrong record to be deleted, appended, or replaced, the program automatically prints the old version so that correction of such an error can be made more readily.

17

The changes listed above take place sequentially in a single pass of the file tape, and a different file tape (a physically different reel on a different tape handler) results. The records of the new file will be numbered sequentially (1,2,...). Thus many, if not all, records will have new record numbers. These new record numbers must be used for the next updating phase (the next month, say).

In any updating phase the changes are assumed to be arranged in ascending order by record number. If several changes of different kinds refer to the same record number, say to 1017, the following sequence of change types (each with its own extra cards) is suggested:

Col 1	Col 21	
INSERTbBEFORE	1017	} These first
INSERTbBEFORE	1017	
REPLACEbbbAUTHOR	1017	
REPLACEbbbLINE	1017	} These next (or REPLACE 1017
APPEND	1017	or DELETE 1017, if appropriate)
APPEND	1017	
INSERTbAFTER	1017	} These last
INSERTbAFTER	1017	

If a change cannot be carried out (due to human missequencing, for example), an error message will be printed to identify the change and the error involved.

The cards of the very last change desired should be followed by a sentinel card with commas in columns 1-5 only. This signals the program to copy any remaining tape records unchanged and terminate the output tape. (See Figure 7.)

No new entries are being made for List A, now published, but a new List B has been started. APPENDS to the List A tape are occasionally added for completeness, however.

```
Col                    Col
1                      21

UPDATE
INSERTbBEFORE          79
:BROWN,bA.C.
THEbHEATbEQUATION.
PHYS.bREV.b28(1968),b128-133.bMRb22,bNO.b9614.
APPEND                 157
MRb23,bNO.bA3929.
DELETE                 271
REPLACEbbbAUTHOR       1079
:MAKLIN,bA.R.
,,,,,
```

Figure 7. Sample Update Specifications

A variation of UPDATE called MERGE can be used to combine two file
tapes of similar form such as List A and List B into a single file. The
distinction between MERGE and UPDATE is that in MERGE the records being
inserted are on the second tape rather than on cards, whereas the
specifications (INSERTbAFTER, etc.) are on cards. When List B is
completed, indexed, and published, it will probably be merged into
List A for easier future search, but publication of the combined list
will be unnecessary. A new List C will probably be started also.

PAGING FOR PUBLICATION

Appropriate use of program options allows page editing for
publication. Record lines do not exceed 87 characters, but some lines
are of this length. Straight listing on the BRL printers at 10
characters per inch would require just under 9 inches plus side margins.
This is in excess of standard publication width; so photographic
reduction by some factor is required.

Reduction of the standard 10 inch length page by the same factor
would result in much wasted space on each page and hence many extra
pages. To avoid this, one can use longer tabulator pages with appropriate
page ejections at the proper point. Page numbers can be inserted (if
known) at the same time by the editing program.

19

This was done in preparing List A for publication. It was determined that the first page of the reference file itself was to be page number 7 and that a full page should contain 68 lines followed by a blank line, the page number line (inserted by the computer), and an "eject page" signal. These parameters were incorporated into an option in the UPDATE mode with a dummy request to modify a non-existent record beyond the end of the file. The output of this computer run was a magnetic tape in properly paged format. Listing on the IBM 1403 printer with longer than normal paper gave a printed file ready for photographic reduction and publication.

Similar paging was applied to the tape of selected index terms and related record numbers. The index was then ready for publication also in the same report.

CONCLUSION

One volume of the file (List A) has been published by BRL and can be obtained for reference. It also contains an index of some 150 significant index terms and the related record numbers, with some cross-referencing. Computer search is recommended, however, for researchers with different requirements.

Both List A and its successor (List B) are available on magnetic tape for search by the BRLESC 1 computer.

REFERENCES

1. John H. Giese, "A Bibliography for the Numerical Solution of Partial Differential Equations," Ballistic Research Laboratories Memorandum Report No. 1991, July 1969, AD 692 827.

2. L. W. Campbell and G. A. Beck, "The FORAST Programming Language for ORDVAC and BRLESC (Revised)," Ballistic Research Laboratories Report No. 1273, March 1965.

3. L. W. Campbell and G. A. Beck, "BRLESC FORTRAN IV," Ballistic Research Laboratories Report No. 1346, October 1966.

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
U.S.Army Aberdeen Research and Development Center Ballistic Research Laboratories Aberdeen Proving Ground, Maryland	Unclassified
	2b. GROUP

3. REPORT TITLE

A COMPUTER BASED SEARCHABLE FILE OF JOURNAL REFERENCES IN THE FIELD OF PARTIAL DIFFERENTIAL EQUATIONS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

George C. Francis

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
February 1970	26	3
8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)	
b. PROJECT NO. RDT&E 1T061102A14B	BRL Memorandum Report No. 2025	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
d.		

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	U.S. Army Materiel Command Washington, D.C.

13. ABSTRACT

The BRLESC computer has been programmed to prepare a searchable and updatable magnetic tape file of journal references in the field of partial differential equations. Multiple simultaneous queries involving Boolean logic are permitted and have been used in automatic indexing f the file prior to publication of a first version. Search strategies are partially self-organizing. Examples of program use are included in the report.

DD FORM **1473** REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

Unclassified
Security Classification

KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Information retrieval						
Automatic indexing						
Journal reference file						
File maintenance						
Prepublication editing						
Boolean logic						
Multiple query search						
Self-organizing search						
Magnetic tape search						
Partial differential equations						
Numerical methods						
BRLESC computer program						

Unclassified
Security Classification